7 Nichtlineare Optimierung

Bisherige Optimierungsmodelle

min Z(**x**)
udN **x** M

charakterisiert durch
- lineare Zielfunktion
- Beschränkung zulässigen Bereichs
  - per Wertebereich Entscheidungsvariablen (zB **x** $\mathbf{R}_+^n$)
  - per Nebenbedingungen, lineare Funktionen in **x**

Nichtlineare Optimierungsmodelle liegen vor, wenn
- Zielfunktion nichtlinear
- und / oder Nebenbedingungen nichtlinear

Entscheidungsvariablen durchgängig reellwertig
- **x** $\mathbf{R}^n$ **unrestringiertes** Problem
- **x** M $\mathbf{R}^n$ **restringiertes** Problem, Restriktionen als Nebenbedingungen-Ungleichungssystem

Generell

- nichtlineare Nebenbedingungen "schwieriger" als nichtlineare Zielfunktion

- restringierte Probleme "schwieriger" als unrestringierte Probleme

- zwischen lokalen / relativen und globalen / absoluten Minima (Maxima) zu unterscheiden

- Hoffnung auf "generelle Methode" (angesichts der Vielfalt nichtlinearer Funktionen) offensichtlich müßig

7.1 Grundbegriffe und Optimalitätsbedingungen

Optimierungsmodell

min Z(**x**) — **x** n-dimensional, Z auf $\mathbf{R}^n$ definiert
udN **x** M — ≠ M $\mathbf{R}^n$

- **x** M → **x** ist **zulässige Lösung**
- Z(**x***) ≤ Z(**x**) **x** M → **x*** ist **globaler Minimalpunkt**
  Z* := min $_{\mathbf{x}\ M}$ Z(**x**) → Z* ist **globales Minimum**
- Satz von **Weierstraß**:
  M kompakt, Z stetig auf M
  globaler Minimalpunkt von Z auf M existiert
- ε > 0
  U_ε(**x**) := {**x'** $\mathbf{R}^n$; |**x**-**x'**| < ε}
  → U_ε(**x**) ist (offene) ε-**Umgebung** von **x**
- ε> 0: Z(**x**) ≤ Z(**x'**) **x'** M∩U_ε(**x**)
  → **x** ist **lokaler Minimalpunkt**
  Z(**x**) ist **lokales Minimum**

Für lineare Optimierungsmodelle war
- lokaler Minimalpunkt auch globaler Minimalpunkt
- M konvexes Polyeder (mit endlich vielen Ecken),
  +: Existenz optimaler Lösung Lösung in Eckpunkt

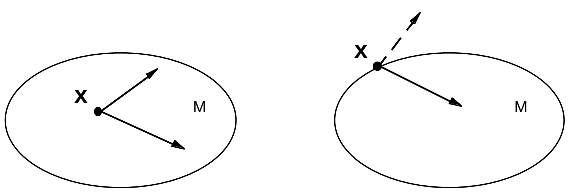
Für nichtlineare Optimierungsmodelle gilt keine dieser Eigenschaften im allgemeinen

LOKALES VORANSCHREITEN

Simplex bewegte sich
von zulässigem Punkt (Eck)
über zulässige Punkte (entlang Kante)
in Z-verbessernder Richtung (Z monoton fallend)
zu zulässigem Punkt (Eck)

Versuch der Übertragung der Idee benötigt Fassungen für
- "über zulässige Punkte"
- "in Z-verbessernder Richtung"
- uam

- **x** M, **s** $\mathbf{R}^n$, δ>0: (**x** + δ' **s**) M 0≤δ'≤δ
  → **s** ist **zulässige Richtung**
  x im Innern von M: alle **s** $\mathbf{R}^n$ zulässig
  x auf Rand von M: nicht alle **s** $\mathbf{R}^n$ zulässig



- Z partiell differenzierbar nach **x**, **x** $\mathbf{R}^n$
  g(**x**) := grad Z(**x**) := (∂Z(**x**)/∂x_1,…,∂Z(**x**)/∂x_n)^T
  → **g** ist **Gradient** von Z in **x**
  weist in Richtung stärksten Anstiegs Z,
  Z(**x**) = z → **g** orthogonal zu Hyperfläche Z(**x**) ≡ z
  → - **g** ist Richtung größter Verbesserung (Verkleinerung) von Z, -**g** orthogonal zu Z(**x**) ≡ z

Bezeichne
S(**x**) := { **s** ; **s** zulässige Richtung in **x** }
die Menge aller in **x** zulässigen Richtungen

Satz 7.1.01 : Verbessernde Richtungen

Sei Z(**x**) stetig differenzierbar
x, **s** $\mathbf{R}^n$
$\mathbf{s}^T$ **g**(**x**) < 0 ($\mathbf{s}^T$ (-**g**(**x**)) > 0)

δ> 0: Z(**x** + δ' **s**) < Z(**x**) 0≤δ'≤δ

oB

Veranschaulichung:
wenn $\mathbf{s}^T$ **g**(**x**) < 0, dann bilden **s** und **g** stumpfen Winkel
s und -**g** spitzen Winkel
Fortschreiten im spitzen Winkel zu -**g** verkleinert Z
x M, **s** S(**x**) zulässiges, verbesserndes Voranschreiten

OPTIMALITÄTSBEDINGUNGEN

Satz 7.1.02 : Notwendige Bedingungen

Sei Z(**x**) stetig differenzierbar
x lokaler Minimalpunkt von Z auf M

$\mathbf{s}^T$ **g**(**x**) ≥ 0 **s** S(**x**) (1)
x innerer Punkt von M **g**(**x**) = 0 (2)

entsprechend Def's lokaler Minimalpunkt
+ zulässige Richtung:
In einer Umgebung von **x** kann Z nicht verkleinert werden
→ **x** ist **stationärer Punkt**
x innerer Punkt von M S(**x**)=$\mathbf{R}^n$ (2)

Für zweimal partiell differenzierbares Z(**x**)
heißt (symmetrische) Matrix
der zweiten partiellen Ableitungen **Hesse-Matrix H**

$$\mathbf{H} := \begin{pmatrix} \frac{\partial^2 Z}{\partial x_1^2} & \cdots & \frac{\partial^2 Z}{\partial x_1 \partial x_n} \\ \cdots & & \cdots \\ \frac{\partial^2 Z}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 Z}{\partial x_n^2} \end{pmatrix}$$

und **H**(**x**) Hesse-Matrix am Punkt **x**

symmetrische n x n Matrix **A** heißt

positiv semidefinit wenn $\mathbf{x} \; \mathbf{R}^n, \mathbf{x} \neq 0$: $\mathbf{x}^T \mathbf{A} \mathbf{x} \geq 0$

positiv definit wenn $\mathbf{x} \; \mathbf{R}^n, \mathbf{x} \neq 0$: $\mathbf{x}^T \mathbf{A} \mathbf{x} > 0$

($\mathbf{x}^T \mathbf{A} \mathbf{x}$ ist **quadratische Form**)

Beispiel:

$$\mathbf{x}^T \mathbf{A}\, \mathbf{x} = (x_1, x_2) \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$
$$= (x_1 - x_2, -x_1 + x_2) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$
$$= (x_1 - x_2) x_1 + (-x_1 + x_2) x_2$$
$$= (x_1 - x_2)^2 \quad \mathbf{x} \neq \mathbf{0}: \; > 0$$

Es wird im folgenden
auf positive Definitheit / Semidefinitheit von **H** ankommen

Daher Hinweise auf Feststellung dieser Eigenschaft(en)

n x n Matrix **A** läßt sich (zB mithilfe Gauß-Algorithmus)
als Produkt **A** = **L R** von Dreiecksmatrizen **L**, **R** darstellen
mit unterer / oberer Dreiecksform

$$\mathbf{L} = \begin{pmatrix} 1 & & & 0 \\ l_{21} & 1 & & \\ \vdots & & \ddots & \\ l_{n1} & \cdots & l_{n,n-1} & 1 \end{pmatrix} \qquad \mathbf{R} = \begin{pmatrix} r_{11} & \cdots & r_{1n} \\ & \ddots & \vdots \\ 0 & & r_{nn} \end{pmatrix}$$

Für symmetrisches **A** ferner **R** als Produkt $\mathbf{R} = \mathbf{D}\,\mathbf{L}^T$
mit Diagonalmatrix
$\mathbf{D} = \mathrm{diag}(d_{11}, \ldots, d_{nn})$

Aus	$\mathbf{A} = \mathbf{L}\,\mathbf{D}\,\mathbf{L}^T$
folgt	$\mathbf{x}^T \mathbf{A}\,\mathbf{x} = \mathbf{x}^T \mathbf{L}\,\mathbf{D}\,\mathbf{L}^T \mathbf{x}$
bzw mit	$\mathbf{y} := \mathbf{L}^T \mathbf{x}$
auch	$\mathbf{x}^T \mathbf{A}\,\mathbf{x} = \mathbf{y}^T \mathbf{D}\,\mathbf{y} = \Sigma\, d_{ii}\, y_i^2$

A genau dann positiv definit, wenn alle $d_{ii} > 0$
positiv semidefinit, wenn alle $d_{ii} \geq 0$

Zielführend folgender Satz

Satz 7.1.03 : Hinreichende Bedingungen

Sei Z(**x**) zweimal stetig differenzierbar
x innerer Punkt von M
g(**x**) = **0**
H(**x**) positiv definit

x ist lokaler Minimalpunkt von Z auf M

oB

Hinweis: positiv semidefinit ist nur notwendige Bedingung
(Analogon im Eindimensionalen: "Wendepunkt")

Nach Satz 7.1.02 ist
$\mathbf{s}^T \mathbf{g}(\mathbf{x}) \geq 0 \qquad \mathbf{s} \quad S(\mathbf{x})$
notwendige Optimalitätsbedingung

- reduziert sich für unrestringierte Probleme auf
  g(**x**) = 0
- ist für restringierte Probleme
  (wo Optimalpunkte auf Rand von M liegen können / werden)
  praktisch schlecht handhabbar
  (wegen Schwierigkeit der expliziten Notierung von S(**x**))

Methode der Lagrange'schen Multiplikatoren
- Integration Nebenbedingungen in Zielfunktion
- Optimierung unrestringierten Problems

bietet Abhilfe,

ist
- von Nebenbedingungs-Gleichungen
- auf Nebenbedingungs-Ungleichungen

zu übertragen

Für Optimierungsproblem

(O) min Z(**x**)
udN $z_i(\mathbf{x}) \leq 0$ i=1,…,m

werden **Lagrange-Multipikatoren** u_i (i=1,…,m) eingeführt
und die **Lagrange-Funktion**
L: $\mathbf{R}^{n+m} \to \mathbf{R}$ definiert gemäß

$$L(\mathbf{x},\mathbf{u}) := Z(\mathbf{x}) + \sum_{i=1}^{m} u_i z_i(\mathbf{x}) = Z(\mathbf{x}) + \mathbf{u}^T \mathbf{z}(\mathbf{x})$$

Punkt ($\mathbf{x}^* \; \mathbf{R}^n, \mathbf{u}^* \; \mathbf{R}_+^m$) heißt **Sattelpunkt** von L, wenn
(vgl. Spieltheorie)
$L(\mathbf{x}^*,\mathbf{u}) \leq L(\mathbf{x}^*,\mathbf{u}^*) \leq L(\mathbf{x},\mathbf{u}^*) \qquad \mathbf{x} \; \mathbf{R}^n, \; \mathbf{u} \; \mathbf{R}_+^m$

Satz 7.1.04 : Sattelpunkt und Optimalität

($\mathbf{x}^* \; \mathbf{R}^n, \mathbf{u}^* \; \mathbf{R}_+^m$) **Sattelpunkt** von L
x* ist optimale Lösung von (O)

- $Z(\mathbf{x}^*) + \mathbf{u}^T \mathbf{z}(\mathbf{x}^*) = L(\mathbf{x}^*,\mathbf{u})$
  $\leq L(\mathbf{x}^*,\mathbf{u}^*) = Z(\mathbf{x}^*) + \mathbf{u}^{*T} \mathbf{z}(\mathbf{x}^*)$
  (a) $\mathbf{u}^T \mathbf{z}(\mathbf{x}^*) \leq \mathbf{u}^{*T} \mathbf{z}(\mathbf{x}^*) \qquad \mathbf{u} \; \mathbf{R}_+^m$
- (a) ist für $\mathbf{z}(\mathbf{x}^*) > \mathbf{0}$, $\mathbf{u} > \mathbf{u}^*$ nicht erfüllbar
  $\mathbf{z}(\mathbf{x}^*) \leq \mathbf{0}$ **x*** zulässig für (O)
- (a) liefert für **u** = **0**
  $\mathbf{0} \leq \mathbf{u}^{*T} \mathbf{z}(\mathbf{x}^*)$
  zusammen mit
  $Z(\mathbf{x}^*) + \mathbf{u}^{*T} \mathbf{z}(\mathbf{x}^*) = L(\mathbf{x}^*,\mathbf{u}^*)$
  $\leq L(\mathbf{x},\mathbf{u}^*) = Z(\mathbf{x}) + \mathbf{u}^{*T} \mathbf{z}(\mathbf{x})$
  $Z(\mathbf{x}^*) \leq Z(\mathbf{x}) + \mathbf{u}^{*T} \mathbf{z}(\mathbf{x}) \qquad \mathbf{x} \; \mathbf{R}^n$
- $\mathbf{u} \geq \mathbf{0}$, $\mathbf{z}(\mathbf{x}) \leq \mathbf{0}$ im zulässigen Bereich
  $Z(\mathbf{x}^*) \leq Z(\mathbf{x})$ zulässigen **x**
  x* ist optimal

"Schön" wäre jetzt,
wenn auch optimaler Punkt Sattelpunkt gälte

Ist problem(-klassen)-abhängig zu untersuchen

7.2 Konvexe Optimierungsprobleme

Sätze 7.1.02/.03 formulieren Bedingungen
für lokale Minimalpunkte fest

bei linearen Optimierungsproblemen waren
- lokale Minimalpunkte
- auch globale Minimalpunkte

Kann nicht allgemein gelten
Aber über lineare Zielfunktionen hinaus?

Erinnerung:

- **Menge** K $\mathbf{R}^n$ heißt **konvex** wenn

  $\mathbf{x}_1,\mathbf{x}_2$ K; 0≤r≤1 r $\mathbf{x}_1$ +(1-r)$\mathbf{x}_2$ K

  Durchschnitt konvexer Mengen ist konvex

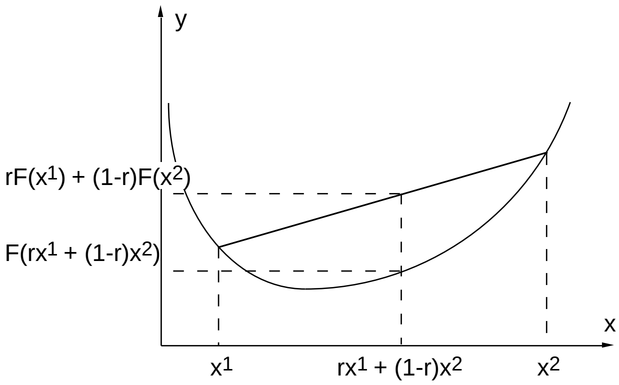
- Sei K $\mathbf{R}^n$ konvex. **Funktion** F: K→**R** heißt **konvex** wenn

  $\mathbf{x}_1,\mathbf{x}_2$ K; 0<r<1 F(r $\mathbf{x}_1$ +(1-r)$\mathbf{x}_2$)
  ≤ r F($\mathbf{x}_1$) +(1-r) F($\mathbf{x}_2$)

  streng konvex für <
  konkav für ≥
  streng konkav für >

- lineare Funktionen sind
  sowohl konvex als auch konkav
  weder streng konvex noch streng konkav

- für K=**R**:
  streng konvexe Funktionen "hängen nach unten durch"



Direkt aus Definitionen:

Satz 7.2.01 : Eigenschaften konvexer Funktionen

Seien K $\mathbf{R}^n$ konvexe Menge
$F_1,\dots,F_m$: K→**R** konvexe Funktionen
a **R**
dann
(a) nichtnegative Linearkombination der F_i

$$G := \sum_{i=1}^{m} r_i F_i \quad r_i \geq 0$$

ist konvexe Funktion auf K

(b) Mengen
{**x** K| F(**x**)≤a} (abgeschlossen)
{**x** K| F(**x**)<a} (offen)
sind konvexe Mengen

Es gibt diverse nützliche Kriterien zur Feststellung
der Konvexität einer Funktion, so

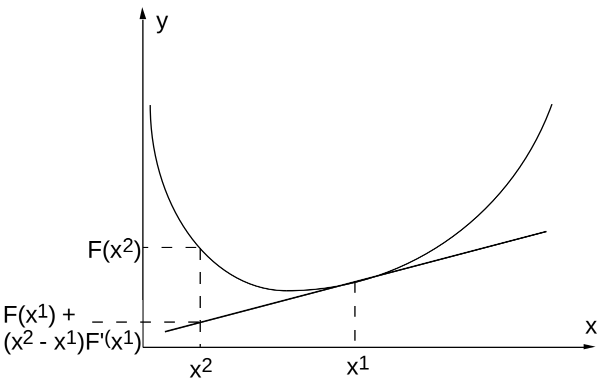
Satz 7.2.02 : Konvexitätskriterium I

Sei K $\mathbf{R}^n$ konvexe Menge mit inneren Punkten
F: K → **R** zweimal stetig differenzierbar
F ist genau dann konvex (streng konvex) auf K wenn
Hesse-Matrix **H**(**x**) positiv semidefinit (positiv definit)
oB

Satz 7.2.03 : Konvexitätskriterium II

Sei K $\mathbf{R}^n$ konvexe Menge
F: K → **R** stetig differenzierbar
F ist genau dann konvex (streng konvex) auf K wenn
$\mathbf{x}_1,\mathbf{x}_2$ K: F($\mathbf{x}_2$) ≥ F($\mathbf{x}_1$) + ($\mathbf{x}_2$-$\mathbf{x}_1$)^T grad F($\mathbf{x}_1$)
(>) oB

Tangente an F(**x**) "niemals oberhalb Kurve":



Mit konvexer Zielfunktion Z(**x**),
konvexen Nebenbedingungen z_i(**x**) ≤ 0 i=1,…,m
und damit konvexer Lösungsmenge
(Durchschnitt von Mengen gemäß S.7.2.01b)
liegt ein sog. **konvexes Optimierungsproblem** (K) vor

(K) min Z(**x**)
udN z_i(**x**) ≤ 0 i=1,…,m
Z, z_i (i=1,…,m) konvex

für konvexe Optimierungsprobleme gilt (erwünschterweise:)

Satz 7.2.04 : Hauptsatz der konvexen Optimierung

Sei M $\mathbf{R}^n$ konvexe (Lösungs-)Menge, M≠
Z: M→**R** konvexe (Ziel-)Funktion
dann
(a) Menge M* aller globalen Minimalpunkte
von Z auf M ist konvex

(b) jeder lokale Minimalpunkt von Z auf M
ist auch globaler Minimalpunkt

zu (a): **x*** globaler Minimalpunkt mit Minimum Z* := Z(**x***)
S.7.2.01b: {**x** M| F(**x**)≤Z*} ist konvex

zu (b): **x**' lokaler Minimalpunkt mit Minimum Z' := Z(**x**')
Annahme: **x** : Z(**x**) < Z' (**x**' nicht global. Min.Pkt)
Z konvex: für 0<r<1 gilt
Z(r **x** +(1-r)**x**') ≤ r Z(**x**) +(1-r) Z(**x**')
Z(**x**' + r(**x**-**x**')) < Z'
in ε-Umgebung von **x**' finden sich **x**'' : Z(**x**'') < Z'
Widerspruch

Insbesondere werden

- die notwendigen Bedingungen für lokalen Minimalpunkt
  (B) Satz 7.1.02: $\mathbf{s}^T \mathbf{g}(\mathbf{x}) \geq 0$ **s** S(**x**)
  $\mathbf{g}(\mathbf{x}) = \mathbf{0}$ für innere Punkte v. M
- für konvexe Optimierungsprobleme
  zu hinreichenden Bedingungen für globalen Minimalpunkt

Satz 7.2.05 : Hinreichende Optimalitätsbedingung

Sei M $\mathbf{R}^n$ konvexe (Lösungs-)Menge
Z: M → **R** konvexe, stetig differenzierb. (Ziel-)Funktion
x* M
(B) gilt für **x*** **x*** ist globaler Minimalpunkt von Z auf M

x ≠ **x***, **x** M
M konvex: Verbindungsstrecke **x*** → **x** in M
x-**x*** S(**x***) (ist zulässige Richtung)
s S(**x***): δ> 0: δ **s** = **x**-**x***
(B) gilt: $(\mathbf{x}-\mathbf{x}^*)^T \mathbf{g}(\mathbf{x}) \geq 0$ **x** M
S. 7.2.03: $F(\mathbf{x}) \geq F(\mathbf{x}^*) + (\mathbf{x}-\mathbf{x}^*)^T \mathbf{g}(\mathbf{x}) \geq F(\mathbf{x}^*)$ **x** M
Behauptung

Auf Umkehrung von S. 7.1.04 (Sattelpunkt Minimalpunkt)
für konvexe Probleme zielend,
Formulierung der sog. **Slater-Bedingung**:

x $\mathbf{R}^n$: nichtlin. Nebenbedingungen: $z_i(\mathbf{x})<0$

(bei ausschließlich nichtlin. Nebenbedingungen
impliziert Slater-Bedingung, daß M innere Punkte enthält)

Satz 7.2.06 : Satz von Kuhn / Tucker

Sei im Rahmen eines konvexen Optimierungsproblems
gemäß (K) die Slater-Bedingung erfüllt.
Dann gilt: **x*** ist optimale Lösungs von (K)
⇔ Lagrange-Funktion L(**x**,**u**) besitzt
Sattelpunkt (**x***,**u***) mit **u*** ≥ **0**
oB

Globale Sattelpunkteigenschaft
ia schlecht explizit nachprüfbar
Für (wie hier gegeben)
differenzierbare, konvexe Z, z_i
durch äquivalente **lokale** Bedingungen ersetzbar:

Karush-Kuhn-Tucker-Bedingungen
für stetig differenzierbare, konvexe Z, z_i (i=1,…,m)

(aus: $\partial L/\partial \mathbf{x}=\mathbf{0}$, $\partial L/\partial \mathbf{u}=\mathbf{0}$
unter Nutzung $\mathbf{z}(\mathbf{x}) \leq \mathbf{0}$, $\mathbf{u} \geq \mathbf{0}$)

(7.2.07) $\operatorname{grad} Z(\mathbf{x}^*) + \sum_{i=1}^{m} u^*_i \operatorname{grad} z_i(\mathbf{x}^*) = \mathbf{0}$

$u^*_i\, z_i(\mathbf{x}^*) = \mathbf{0}$ i = 1,…,m
$z_i(\mathbf{x}^*) \leq \mathbf{0}$ i = 1,…,m
$\mathbf{u}^* \geq \mathbf{0}$

KKT-Bedingungen

- lassen sich für spezifische (Unter-)Klassen
  weiter konkretisieren
- so zB (breit untersucht) für **quadratische Opt.Probleme**
  "=" quadratische Zielfunktion
  \+ lineare Nebenbedingungen (M ist konv. Poyeder)

7.3 Übersicht über Lösungsverfahren

Skizzen zu Fällen
- EINDIMENSIONALE OPTIMIERUNG
- UNRESTRINGIERTE OPTIMIERUNGSPROBLEME IM $\mathbf{R}^n$
- RESTRINGIERTE OPTIMIERUNGSPROBLEME IM $\mathbf{R}^n$

EINDIMENSIONALE OPTIMIERUNG

Z: **R** → **R** Zielfunktion

min Z(x)
udN x **R** bzw x $\mathbf{R}_+$ bzw x [a,b], b>a

- einfachste Optimierungsaufgabe mit eigenem Interesse
- Teilproblem in höherdimensionalen Optimierungsaufgaben

Problem besitzt genau eine optimale Lösung,
wenn Z im zulässigen Bereich [a,b]
unimodal x* [a,b]: Z streng monoton fallend auf [a,x*],
Z streng monoton steigend auf [x*,b]
streng konvex unimodal

Einfache Lösungsverfahren ("ableitungsfrei")
- nutzen nur Funktionswerte
- verkleinern "Suchintervall" schrittweise

Typisch: 2 "Stützstellen" x_1,x_2 mit $a<x_1<x_2<b$,
$Z(x_1) \leq Z(x_2)$: weiter mit $[a,x_2]$
$Z(x_1) > Z(x_2)$: weiter mit $[x_1,b]$
Teilung zB "goldener Schnitt"

Aufwendigere (idR bessere) Lösungsverfahren
- nutzen Funktionswerte + Werte 1. (+höherer) Ableitungen
- konstruieren Folge (hoffentlich:) konvergierender Punkte

Typisch: **Newton**-Verfahren

- Annahme: Z zweimal stetig differenzierbar
  Minimalpunkt x* (a,b)
  Näherungspunkt x^i (a,b); i=0,1,…
  (Vorbereitung durch "einfaches" Verfahren,
  Grenzpunkte a,b immer separat zu prüfen)
- Ziel: Bestimmung x*:Z'(x*)=0 Nullstelle Z'
  (Z konvex: notwendig und hinreichend
  für globale Minimalstelle)
- Vorgang: Berechnung $Z'(x^i)$, $Z''(x^i)$
  Linearisierung der Ableitungsfunktion Z' in x^i
  Bestimmung der Tangentennullstelle =: x^{i+1}

  $$x^{i+1} := x^i - \frac{Z'(x^i)}{Z''(x^i)}$$
- Newton verwendet quadratische (parabolische)
  Approximation von Z
  bei quadratischem Z Optimalstelle in 1 Schritt gefunden

  Konvergenz bei beliebigem Startpunkt x^0 nicht gesichert,
  analytische Bedingungen für sicheres x^0 kompliziert
- sind Z', Z'' nicht analytisch bekannt,
  kann mit numerischen Näherungen
  auf Basis der Funktionswerte "benachbarter" Punkte
  gearbeitet werden

OPTIMIERUNGSPROBLEME IM $\mathbf{R}^n$

- generell iterativ: $\mathbf{x}^0 \rightarrow \mathbf{x}^1 \rightarrow \ldots \rightarrow \mathbf{x}^i \rightarrow \ldots$
  im zulässigen Bereich: $\mathbf{x}^i$ M $\mathbf{R}^n$, i=0,1,…
  idR nicht inhärent abbrechend Näherungslösungen
- (Hoffnung auf) Konvergenz zu optimalem Punkt $\mathbf{x}^*$:

  $\mathbf{s}^T \mathbf{g}(\mathbf{x}^*) \geq 0$ **s** S($\mathbf{x}^*$)
  bzw $\mathbf{g}(\mathbf{x}^*) = \mathbf{0}$ im unrestringierten Fall: M = $\mathbf{R}^n$

  bei nichtkonvexem Problem lediglich stationärer Punkt
  nicht notwendig Extrempunkt, optimaler Punkt
  bezeichne LM die Lösungsmenge (stationäre Punkte)
  LM := { $\mathbf{x}^*$ M | $\mathbf{s}^T \mathbf{g}(\mathbf{x}^*) \geq 0$ **s** S($\mathbf{x}^*$) }
  bzw LM := { $\mathbf{x}^*$ M | $\mathbf{g}(\mathbf{x}^*) = \mathbf{0}$ }
- Konvergenz, Konvergenzgeschwindigkeit
  im konkreten Fall zu prüfen

UNRESTRINGIERTE OPTIMIERUNGSPROBLEME IM $\mathbf{R}^n$

min Z($\mathbf{x}$) (mindestens 1-mal) stetig differenzierbar
udN x $\mathbf{R}^n$

LM = { $\mathbf{x}^*$ M | $\mathbf{g}(\mathbf{x}^*) = \mathbf{0}$ }

Behandlungsmöglichkeit mit **Abstiegsverfahren**

Schritte von Abstiegsverfahren

- Initialisierung
  wähle $\mathbf{x}^0$ $\mathbf{R}^n$, i:=0
- Iteration über
  - Abbruchprüfung
    berechne $\mathbf{g}^i := \mathbf{g}(\mathbf{x}^i)$
    terminiere falls $\mathbf{g}^i = \mathbf{0}$ (bzw $|\mathbf{g}^i| < \varepsilon$)
  - Abstiegsrichtung, Suchrichtung
    bestimme $\mathbf{s}^i$ $\mathbf{R}^n$: $(\mathbf{s}^i)^T \mathbf{g}^i < 0$
  - Schrittweite
    berechne $d^i > 0$: $Z(\mathbf{x}^i + d^i \mathbf{s}^i) < Z(\mathbf{x}^i)$
    optimale Schrittweite ist eindimensionales Subproblem
  - Fortschritt
    $\mathbf{x}^{i+1} := \mathbf{x}^i + d^i \mathbf{s}^i$; i := i+1

Konkrete Verfahren unterscheiden sich durch Wahl von
- Abstiegsrichtung
- Schrittweite

- klassisches **Gradientenverfahren**, **Verfahren des steilsten Abstiegs**

  $\mathbf{s}^i := -\mathbf{g}^i$
- idR deutlich schneller konvergierend ist mehrdimensionales **Newtonverfahren**
  - zusätzlich benötigt: zweite Ableitungen
    Forderung Z (mindest.) 2-mal stetig diff'bar
    Notation $\mathbf{H}^i := \mathbf{H}(\mathbf{x}^i)$ Hesse-Matrix

$\mathbf{H}^i$ positiv definit, Satz 7.2.02
Z ist in Umgebung von $\mathbf{x}^i$ streng konvex

- Standard-Newton setzt
  $\mathbf{s}^i := -(\mathbf{H}^i)^{-1}\mathbf{g}^i$
  $d^i := 1$ (ia nicht optimal, optimieren!)
  so daß
  $\mathbf{x}^{i+1} := \mathbf{x}^i - (\mathbf{H}^i)^{-1}\mathbf{g}^i$
- $\mathbf{H}^i$ muß nicht explizit invertiert werden, da wegen
  $\mathbf{H}^i \mathbf{x}^{i+1} := \mathbf{H}^i \mathbf{x}^i - \mathbf{g}^i$
  $\mathbf{x}^{i+1}$ als Lösung linearen Gl'systems ermittelbar
- Nachteil ist hoher Rechenaufwand:
  Hesse-Matrix + lin. Gl'system
  abgemildert in vereinfachtem Newton-Verfahren,
  welches $\mathbf{H}^i$ für "einige" Schritte i,i+1, … verwendet

- "zwischen"
  - Gradientenverfahren (schlechte Konvergenz)
  - Newtonverfahren (hoher Aufwand)

  sind **Verfahren der konjugierten Gradienten**
  - mit besserer Konvergenz als Gradientenverfahren
  - mit geringerem Aufwand als Newtonverfahren

  …
- …

RESTRINGIERTE OPTIMIERUNGSPROBLEME IM $\mathbf{R}^n$

min Z($\mathbf{x}$) idR stetig differenzierbar
udN **x** M $\mathbf{R}^n$

LM := { $\mathbf{x}^*$ M | $\mathbf{s}^T \mathbf{g}(\mathbf{x}^*) \geq 0$ **s** S($\mathbf{x}^*$) }

(Hoffnung auf) Näherung für stationäre Punkte $\mathbf{x}^*$ LM
Z, M konvex (Satz 7.2.05) $\mathbf{x}^*$ sind glob. Minimalpunkte

Verfahrensklassen:
- Verfahren zulässiger Richtungen
- Verfahren der Straffunktionen (Barrierefunktionen)
- Schnittebenenverfahren
- Karush-Kuhn-Tucker-Verfahren
- …

- **Verfahren zulässiger Richtungen**

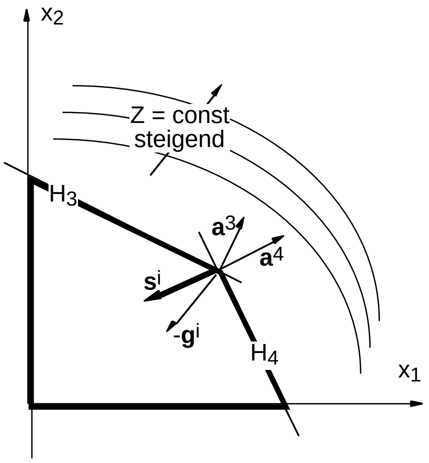
Anwendung iW für lineare Restriktionen

Vorgang in Abwandlung von nichtrestringiertem Vorgang

- Initialisierung: Wahl $\mathbf{x}^0$ M, i:=0
- Iteration über
  * Abstiegsrichtung: Bestimmung $\mathbf{s}^i$ $\mathbf{R}^n$: $(\mathbf{s}^i)^T \mathbf{g}^i < 0$
    (ohne Erfolg: Abbruch)
  * Schrittweite: Berechnung $d^i > 0$: $Z(\mathbf{x}^i + d^i \mathbf{s}^i) < Z(\mathbf{x}^i)$
    mit optimaler Schrittweite,
    aber **unter Beachtung Restriktionen**
  * Fortschritt: $\mathbf{x}^{i+1} := \mathbf{x}^i + d^i \mathbf{s}^i$; i := i+1

Bei linearen Restriktionen Nebenbeding'gen gewohnter Form:
min $Z(\mathbf{x})$
udN $\mathbf{A}\,\mathbf{x} \le \mathbf{b}$ $\mathbf{b}$ $\mathbf{R}^m$, **A:** m x n

Iterationspunkt $\mathbf{x}^i$ M (immer gewahrt) liegt
- auf (0,1,…,m) begrenzenden Hyperebenen
  H_k: $(\mathbf{a}^k)^T\,\mathbf{x}^i = b^k$ k I^i
  wo $\mathbf{a}^k$ k-ter Zeilenvektor **A** (Vektor $\mathbf{a}^k$ orth. H_k,
dh "nach außen")
- auf keiner Grenzebene (im Inneren von M) für I^i=
- oder auf Menge **in $\mathbf{x}^i$ aktiver Hyperebenen**



Bedingungen Fortschrittsrichtung $\mathbf{s}^i$:
- im stumpfen Winkel zu $\mathbf{a}^k$, k I^i "zulässig"
- im spitzen Winkel zu $-\mathbf{g}^i$ "Z besser (kleiner)"

Bedingungen formal:

- stumpfe, äußerstenfalls rechte Winkel
  $(\mathbf{a}^k)^T\,\mathbf{s}^i \le 0$ k I^i
- spitzer Winkel $-\mathbf{g}^i$ → stumpfer Winkel $\mathbf{g}^i$
  $(-\mathbf{g}^i)^T\,\mathbf{s}^i > 0$ → $(\mathbf{g}^i)^T\,\mathbf{s}^i < 0$

Verschiedene Regeln zur Wahl von $\mathbf{s}^i$

- Rosen-Verfahren
  * berechnet $\mathbf{s}^i$ als Projektion $\mathbf{g}^i$ auf $\cap(H_k; k\ \ I^i)$ für $I^i \neq$
    nutzt $-\mathbf{g}^i$ als $\mathbf{s}^i$ für I^i=
  * in beiden Fällen ist $|\mathbf{s}^i| < \varepsilon$ Abbruchbedingung
    (stationärer Punkt erreicht)
- vgl auch Karmarkars Projektionsmethode?
  blieb "im Inneren" !

Zusätzlich zu beachten:

- Schrittweite "d^i" so daß M nicht verlassen
  (wegen linearer NB leicht formalisierbar)

  uU zu optimieren
  (eindimensionale Optimierung)
- Begrenzung Schrittweite führt implizit
  auch zur Entdeckung "unbegrenzten Problems"

- **Verfahren der Straffunktionen (Barrierefunktionen)**

Restringiertes Optimierungsproblem

- durch unrestringiertes Optimierungsprobleme approximiert
- unter Veränderung der Zielfunktion derart daß
  * Verlassen des zulässigen Bereichs
    mit "Strafkosten" belastet
    (Veränderung Z außerhalb M)
  * Rand des zulässigen Bereichs
    in "Barriere" verwandelt
    (Veränderung Z bei Annäherung an Rand M)

Skizze **Straffunktionen**:

min $Z(\mathbf{x})$ udN $\mathbf{z}(\mathbf{x}) \le \mathbf{0}$
→ min $Y(\mathbf{x})$ udN $\mathbf{x}$ $\mathbf{R}^n$

wo $Y(\mathbf{x}) := Z(\mathbf{x}) + p(\mathbf{x})$ mit $p(\mathbf{x})$ $= 0$ $\mathbf{x}$ M
> 0 $\mathbf{x}$ M
Minimalpunkt $\mathbf{x}$ M für Y ist Minimalpunkt für Z

verschiedene Straffunktionen, oft verwendet:

$p(\mathbf{x}) := \mathbf{z}^+(\mathbf{x})^T\,\mathbf{z}^+(\mathbf{x})$
wo $\mathbf{z}^+(\mathbf{x}) := \max(\mathbf{0}, \mathbf{z}(\mathbf{x}))$ (komponentenweise)

Folgt Behandlung à la unrestringiertes Problem,
Idee idR verfeinert in Form
$Y(\mathbf{x}) := Z(\mathbf{x}) + r^i\,p(\mathbf{x})$

dh mit
- gewichteter Strafe
+ im Iterationsverlauf angepaßter Gewichtung
  (idR r^i wachsend mit Iterationsindex)

Skizze **Barrierefunktionen**:

Voraussetzung: M besitzt innere Punkte M° M

Barrierefunktionen b: M° → **R** besitzen die Eigenschaften
- b stetig auf M°
- $b(\mathbf{x}) \to \infty$ bei Annäherung von **x** an den Rand von M

verschiedene Barrierefunktionen,

im Falle von Nebenbedingungen
$\mathbf{z}(\mathbf{x}) < \mathbf{0}$ (Rand **nicht** erlaubt)
oft verwendet sind

$$b_1(\mathbf{x}) := -\sum \ln\left(-z_i(\mathbf{x})\right) \qquad b_2(\mathbf{x}) := -\sum \frac{1}{z_i(\mathbf{x})}$$

Folgt Behandlung à la unrestringiertes Problem,
unter Minimierung der Ersatzfunktion

$Y(\mathbf{x}) := Z(\mathbf{x}) + 1/r^i\,b(\mathbf{x})$

mit r^i>0, Vergrößerung der r^i im Iterationsverlauf

- **Schnittebenenverfahren**

Vorrangig bei nichtlinearen Restriktionen eingesetzt

Zunächst lineare Zielfunktion:

min $\mathbf{c}^T\,\mathbf{x}$
udN $\mathbf{x}$ M $\mathbf{R}^n$ M nichtleer, kompakt, konvex,
kein Polytop (sonst: bekannt)

"Um M" werde konvexes Poytop P_1 M gelegt

Optimierung auf P_1 (zB mit Simplex)
→ Lösung $\mathbf{x}^1$

- $\mathbf{x}^1$ M $\mathbf{x}^1$ ist optimale Lösung
- $\mathbf{x}^1$ M → Bestimmung Hyperebene, welche P_1 beschneidet
  → konvexes Poytop P_2
  wo P_2 M, $\mathbf{x}^1$ P_2

usf

Verschiedene Verfahren zur Konstruktion der Schnittebenen im Gebrauch

vgl auch Gomory-Verfahren

(bis hier angenommene)
Linearität der Zielfunktion ist nicht wesentlich:

min	$Z(\mathbf{x})$	Z nichtlinear, konvex
udN	$\mathbf{x}$ M $\mathbf{R}^n$	M nichtleer, kompakt, konvex

Einführung zusätzlicher Variabler x_{n+1} $\mathbf{R}$
+ zusätzlicher Nebenbedingung $Z(\mathbf{x}) - x_{n+1} \le 0$

Betrachtung Optimierungsproblem

min	x_{n+1}	lineare Zielfunktion
udN	$\mathbf{x}$ M $\mathbf{R}^n$	
	$Z(\mathbf{x}) - x_{n+1} \le 0$	

$\mathbf{x}^*$ genau dann optimale Lösung Ursprungsproblem,
wenn $(\mathbf{x}^*, x^*_{n+1})$ Lösung Ersatzproblem, wo $x^*_{n+1} = Z(\mathbf{x}^*)$

- **Karush-Kuhn-Tucker-Verfahren**

Analytische Verfolgung der KKT-Bedingungen
(hinreichend für Optimalität)
für konkrete Problemklassen
(zB quadratische Zielfunktion,
in Abschn.7.2 erwähnt, nicht explizit ausgeführt)
führt ggf zu einfacher behandelbaren Ersatzproblemen
(zB mit Methoden der linearen Optimierung)

hier "Abbruch",
viel weiteres Wissen vorhanden!

(Lagrange-Dualität,
Separable Optimierung,
Quotientenoptimierung,
konkave Optimierung,
B&B-Verfahren,
Deflations- und Tunnel-Techniken,
nichtdeterministische Verfahren,
evolutionäre Algorithmen
…)

LEER

LEER